Article

Exceptional Children
2015, Vol. 81(4) 443–470

DOI: 10.1177/0014402914563702
ec.sagepub.com
SAGE

Effects of a Multitier Support System on Calculation, Word Problem, and Prealgebraic Performance Among At-Risk Learners

Sarah R. Powell[1], Lynn S. Fuchs[2], Paul T. Cirino[3], Douglas Fuchs[2], Donald L. Compton[2], and Paul C. Changas[4]

Abstract

The focus of the present study was enhancing word problem and calculation achievement in ways that support prealgebraic thinking among second-grade students at risk for mathematics difficulty. Intervention relied on a multitier support system (i.e., responsiveness to intervention, or RTI) in which at-risk students participate in general classroom instruction and receive supplementary small-group tutoring. Participants were 265 students in 110 classrooms in 25 schools. Teachers were randomly assigned to one of three conditions: calculation RTI, word problem RTI, or business-as-usual control. Intervention lasted 17 weeks. Multilevel modeling indicated that calculation RTI improved calculation but not word problem outcomes, word problem RTI enhanced proximal word problem outcomes as well as performance on some calculation outcomes, and word problem RTI provided a stronger route than calculation RTI to prealgebraic knowledge.

Calculation (CAL) and word problems (WPs) are critical forms of mathematics competence for success in the primary grades and through adulthood. A CAL problem is set up for solution. By contrast, a WP requires students to process text to build a problem model and then construct a number sentence for calculating the unknown. Along with this transparent difference in the nature of these tasks, correlational studies suggest that different cognitive abilities underlie CAL and WP skill (e.g., Fuchs, Fuchs, Stuebing, et al., 2008; Fuchs, Geary, et al., 2010; Swanson, 2006).

At the same time, students who have concurrent difficulty in CAL and WPs have more severe deficits in each domain (e.g., Fuchs, Fuchs, Stuebing, et al., 2008), and they experience more pervasive challenges on other mathematics competencies. This includes prealgebraic thinking (Powell & Fuchs, 2014), which is problematic because algebra is frequently required for high school graduation (National Center for Education Statistics, 2008), is associated with successful participation in the U.S. workforce, and represents a gateway to more advanced mathematics achievement (National Mathematics Advisory Panel [NMAP], 2008; RAND Mathematics Study Panel, 2003). Interest has therefore increased in identifying ways to

[1]University of Texas at Austin
[2]Vanderbilt University
[3]University of Houston
[4]Metropolitan Nashville Public Schools

Corresponding Author:
Sarah R. Powell, University of Texas at Austin, 1 University Station D5300, Austin, TX 78712, USA.
E-mail: srpowell@austin.utexas.edu

design early mathematics instruction to create a strong platform for algebraic reasoning (see, e.g., NMAP, 2008).

For these reasons, our focus in the present study was enhancing second-grade CAL and WP achievement in ways that simultaneously support prealgebraic thinking. Our target population was students with strong risk for mathematics difficulty (MD): those with concurrent low performance at the start of second grade on CAL and WPs (i.e., CAL/WP-MD). Our approach to intervention relied on a multitier support system, often referred to as *responsiveness to intervention* (RTI; Vaughn & Fuchs, 2003), in which at-risk students participate in general classroom instruction while receiving supplementary small-group tutoring (Fuchs, Fuchs, Craddock, et al., 2008; Lembke, Garman, Deno, & Stecker, 2010; Mellard, Stern, & Woods, 2011). With such tutoring, schools provide students with a learning environment to remediate foundational skills by explaining mathematical concepts and procedures using more simple and direct language, by providing practice in more systematic and novel ways, and by delivering more immediate corrective feedback when students demonstrate incorrect mathematical thinking. In providing this second layer of support, the hope is that small-group tutoring works synergistically with classroom mathematics instruction to strengthen the learning of students with MD.

Interest has increased in identifying ways to design early mathematics instruction to create a strong platform for algebraic reasoning.

However we were also interested in whether a focus on CAL or WPs provides a stronger route toward prealgebraic thinking for students with CAL/WP-MD. We therefore designed and implemented two separate multilevel mathematics systems, one in CAL and another in WPs, and randomly assigned teachers and their classrooms to WP-RTI, CAL-RTI, or business-as-usual control. In this introduction, we provide background information on prealgebra and its possible connections to CAL and WPs. We then outline prior work related to CAL-RTI or WP-RTI and explain the present study's purpose and hypotheses. Note that a parent research report (Fuchs et al., in press) also focused on transfer from CAL and from WP intervention to prealgebraic thinking. That report considered a larger sample of students who received one or two levels of support, whose mathematics skills spanned the continuum, with difficulty in CALs or WPs or in neither area. That report did not disaggregate outcomes according to difficulty status.

Prealgebraic Thinking and Possible Connections to CALs and WPs

In contrast to arithmetic, algebra involves symbolizing and operating on numerical relationships within mathematical structures. Some subscribe to the position that algebra creates an interference effect with arithmetic, and algebra is therefore developmentally inappropriate for young students (e.g., Balacheff, 2001; Linchevski, 2001; Linchevski & Herscovics, 1996; Sfard, 1991). Yet, because algebraic expressions can be treated procedurally, by substituting numerical values to yield numerical results (Kieran, 1990), understanding of arithmetic principles appears to involve generalizations that are algebraic in nature. This suggests algebra warrants a role in early instruction (Blanton & Kaput, 2005; Carraher & Schliemann, 2002; NMAP, 2008). A third view, which bridges both perspectives, supports a connection between arithmetic and algebraic thinking but only if arithmetic instruction—on whole-number CALs or WPs—is designed to facilitate that transition.

Pillay, Wilss, and Boulton-Lewis's (1998) model of learning underpins the third perspective, and Kieran (1990) relied on Pillay et al. to propose a developmental progression. The first stage is arithmetic competence: the capacity to operate numerically and the understanding of operational laws and relational meaning of the equal sign in standard equations (i.e., both sides of the equal sign have the same value). This provides the foundation for a prealgebraic stage that builds on arithmetic competence by expanding the relational meaning of the equal sign to include nonstandard equations (e.g., 4 = 13 – 9; 4 + 8 = 6 + 6), the concept of unknowns in equations, and the concept of a variable. This middle stage supports development of formal algebraic competence and indicates potential targets for early prealgebraic intervention: understanding of the equal sign (i.e., solving nonstandard equations with one unknown) and the concept of a variable (i.e., completing function tables). It also suggests the need for instruction that supports the transition between arithmetic competence and formal algebraic competence.

Unfortunately, research illustrates how conventional instruction fails to address this need. For example, typical instruction causes most students to misconstrue the equal sign as an operational symbol (Baroody & Ginsburg, 1983; McNeil & Alibali, 2005; Powell, 2012). Students commonly answer problems like 7 + 5 = _ + 3 as 12 (ignoring the operation to the right side of the equal sign) or 15 (adding all known values; Falkner, Levi, & Carpenter, 1999). Such confusion persists into high school (NMAP, 2008) and is associated with difficulty in using algebraic notation to represent WPs (Powell & Fuchs, 2010) and solve linear equations (Alibali, Knuth, Hattikudar, McNeil, & Stephens, 2007; Knuth, Stephens, McNeil, & Alibali, 2006).

It is nevertheless unclear whether a focus on arithmetic CAL or arithmetic WPs is more productive for facilitating the transition from arithmetic to algebraic thinking. Fluency with CAL may reflect a strong foundation in arithmetic operational laws and may support prealgebra by reducing demands on working memory to free up attention for the challenges associated with handling nonstandard equations and variables (Geary et al., 2008). On the other hand, WPs require CAL as well as two forms of symbolic representations (numerals and language), even as WPs reflect understanding of relationships between known and unknown quantities (Geary et al., 2008). WPs may therefore involve greater symbolic complexity than CAL and rely more on the type of mental flexibility, manipulation of symbolic associations, and maintenance of multiple representations (numerical and linguistic) that support prealgebraic thinking (e.g., Kieran, 1992; Sfard & Linchevski, 1994).

Lee, Ng, Bull, Pe, and Ho (2011) and Tolar, Lederberg, and Fletcher (2009) have investigated potential connections with algebra, but these studies are correlational and investigated either CAL or WPs. We located only one report (Fuchs et al., 2012) that contrasted the connection to algebra of WPs versus CAL. Although results suggested WPs have a stronger connection, the design was again correlational. What is required is an experimental study examining whether intervention in CAL or WPs improves students' prealgebraic knowledge. In the present study, we focused on students with CAL/WP-MD—a challenging population for expecting arithmetic intervention to result in superior prealgebraic cognition.

Prior Work Related to CAL or WP Small-Group Tutoring or RTI

Previous studies have assessed the efficacy of small-group tutoring to improve the mathematics performance of students with MD, typically defined on a measure of CAL or a broad-based mathematics achievement test. An instructional model that incorporates explicit instruction with multiple visual representations, heuristics, and student verbalizations contributes to improved mathematics performance in this population of learners (Gersten et al., 2009). Most but not all tutoring studies that incorporate such practices have been associated with positive effects (e.g., Fuchs, Powell, et al., 2008; Jitendra, Hoff, & Beck, 1999; Powell & Fuchs, 2010; Tournaki, Bae, & Kerekes, 2008). For example, Bryant, Bryant, Gersten, Scammacca, Funk, et al. (2008) conducted a regression-discontinuity study with first graders who were identified for tutoring due to low performance on a mathematics progress-monitoring measure. Tutoring focused on number recognition, counting, comparing numbers, place value, and number combinations. At posttest, students scored significantly higher than expected. By contrast, Bryant, Bryant, Gersten, Scammacca, and Chavez (2008) found mixed results; second graders, but not first graders, showed performance gains as a function of similarly designed tutoring. Moreover, at posttest, the second graders did not perform comparably to students without MD. The authors thus concluded that more powerful interventions are needed to realize the necessary improvement for this population of learners.

One strategy for increasing intervention power may reside in a multitier framework, such as RTI (Vaughn & Fuchs, 2003), in which students with CAL/WP-MD receive research-based classroom instruction in combination with well-designed small-group tutoring. Yet, RTI is a relatively new innovation, and insufficient attention has been directed at understanding the potential of mathematics RTI for students with MD. We identified only one randomized control trial focused on a multitier support system in mathematics. Fuchs, Fuchs, Craddock, et al. (2008) randomly assigned third-grade classrooms to receive research-principled WP whole-class instruction or to receive the school's standard WP instruction. Then, within each classroom, at-risk students were identified due to poor WP performance and randomly assigned to receive supplementary small-group tutoring support or to continue in the general education program without such support. At posttest, students who received the combination of research-principled whole-class and small-group WP instruction demonstrated substantially better learning than students in the other three conditions. This led the authors to conclude that two tiers of support are better than one tier for students with MD.

In the present study, we therefore focused on the efficacy of two-tier mathematics RTI systems that incorporate whole-class research-principled general education instruction combined with supplementary small-group instruction. We extended the literature in four ways: (a) by targeting a population not previously studied in prior RTI or tutoring studies, students with CAL/CP-MD; (b) by focusing on second grade, when MD is established but when intervention may be early enough to reverse the trajectory of mathematics development; (c) by contrasting the efficacy of a multitier system focused on CAL skill against an analogous system focused on WP skill while also including a business-as-usual control condition; and (d) by examining whether a focus on CAL or WPs within an RTI provides a stronger route toward prealgebra thinking.

The major purpose of this study was to assess the efficacy of two RTI programs, one focused on CAL and the other on WPs.

Note that our major purpose was to examine effects of the two-tier intervention systems on CALs, WPs, and prealgebraic knowledge. Our major contrast, therefore, was between the two active RTI conditions. At the same time, including a business-as-usual control group controlled for maturation and history effects and permitted conclusions about whether students in one or more of the RTI conditions made more progress than would have occurred without RTI. (See Instructional Framework in Method section for information on the instructional approach taken in each RTI system and its potential connection with prealgebra.)

Study Purpose and Hypotheses

The major purpose of this study was, therefore, to assess the efficacy of two RTI programs, one focused on CAL and the other on WPs, specifically for second-grade students with CAL/WP-MD. We contrasted the efficacy of CAL-RTI and WP-RTI against each other and against the control group on CAL, WPs, and prealgebraic knowledge. CAL intervention provided no instruction on WPs; WP intervention provided no instruction on CAL (students were told to use whatever methods their classroom teachers had taught to calculate answers). The second-grade two-tiered CAL intervention was based in part on previously validated whole-class (Fuchs et al., 1997) and tutoring (Fuchs et al., 2011; Fuchs, Powell, et al., 2009) programs at third grade. The second-grade two-tiered WP intervention was based in part on previously validated whole-class (Fuchs, Fuchs, Craddock, et al., 2008) and tutoring (Fuchs et al., 2011; Fuchs, Powell, et al., 2009) programs at third grade. These programs are referred to, respectively, as *Math Wise* and *Pirate Math*.

Our hypotheses were as follows. First, we expected effects to be specific to the focus of intervention: that CAL-RTI would produce superior CAL outcomes, compared to WP-RTI and to control, and that WP-RTI would result in superior WP outcomes, compared to CAL-RTI and to control. Second, because we designed intervention to facilitate the transition from arithmetic to algebra, we expected CAL-RTI and WP-RTI to improve prealgebraic thinking, with both conditions outperforming the control group. We did, however, expect WP-RTI to produce stronger prealgebraic performance than CAL-RTI, in light of Fuchs et al.'s (2012) correlational study and based on the assumption that arithmetic WPs rely more on the type of mental flexibility, manipulation of symbolic associations, and maintenance of multiple representations that support prealgebraic thinking.

Method

Participants

Participants with CAL/WP-MD were selected from 1,917 students in 127 second-grade classrooms taught by 96 teachers in 25 schools in a metropolitan school district across four cohorts (one per year for 4 years). Before recruitment, we received institutional review board approval. All teachers signed consent forms, and all students had signed consent forms from their parent or guardian. Some teachers participated in more than one study cohort, hence the discrepancy between number of classrooms and teachers. Stratifying by school, teachers (and their classrooms) were randomly assigned to three conditions: ~37.5% to CAL intervention (701 students), ~37.5% to WP intervention (705 students), and ~25% to control (511 students). This maximized research-principled intervention while maintaining a large-enough control group. Once assigned, teachers who participated across cohorts remained in their condition, with ~25% of classrooms in each condition having teachers in multiple cohorts. The present study is a quasi-experiment due to a lack of complete randomness for all teachers (i.e., 26 teachers had participated in a previous cohort and retained those assignments, eight teachers received a condition to which they were not assigned, and two teachers fell into both categories). The parent report (Fuchs et al., in press) indicated results were not substantively altered when only those who were newly and appropriately randomly assigned were evaluated.

The present analysis focused only on CAL/WP-MD students, whom we identified based on whole-class screening with the 1,917 students on CAL and WP measures (see Measures). The benchmark for low performance in this report was scoring <7 on the CAL screening measure, <6 on the WP screening measure, and scoring above the 9th percentile on at least one subtest of the two-subtest Wechsler Abbreviated Scale of Intelligence (WASI; Psychological Corporation, 1999). These benchmarks for low performance were first calculated empirically using data from a pilot study where we identified pretest cutoff scores predictive of low performance on CAL and WP outcome measures. For the current report, we adjusted these benchmarks downward so that only students below the 40th percentile in both areas (i.e., CAL and WP) were included. With these cutoff scores, we identified 265 students from 110 second-grade classrooms taught by 83 teachers in 25 schools. Students were assigned to one of three tutoring conditions: 72 in Cohort 1 (27 in CAL-RTI, 24 in WP-RTI, 21 in control [CON]), 56 in Cohort 2 (21 in CAL-RTI, 19 in WP-RTI, 16 in CON), 77 in Cohort 3 (29 in CAL-RTI, 33 in WP-RTI, 15 in CON), and 60

Table 1. Student Demographics by Tutoring Condition.

	CAL (n = 96)		WP (n = 91)		Control (n = 78)	
Variable	n	(%)	n	(%)	n	(%)
Males	47	(49.0)	54	(59.3)	31	(39.7)
Race						
African American	51	(53.1)	51	(56.0)	41	(52.6)
White	10	(10.4)	14	(15.4)	16	(20.5)
Hispanic	30	(31.3)	20	(22.0)	14	(17.9)
Other	5	(5.2)	6	(6.6)	7	(9.0)
RFL	90	(93.8)	79	(86.6)	69	(88.5)
School-identified disability	8	(8.3)	8	(8.8)	11	(14.1)
ESL	18	(18.8)	15	(16.5)	14	(17.9)
Retained	7	(7.3)	4	(4.4)	6	(11.5)

Note. CAL = calculation intervention; WP = word problem intervention; RFL = reduced-price and/or free lunch; ESL = English as a second language.

in Cohort 4 (19 in CAL-RTI, 15 in WP-RTI, 26 in CON). Screening measures correlated moderately at r = .44 in the entire screened sample, and according to their screening scores, the 265 students composed the lowest 14% of the entire screened sample; the mean scores on the screening measures were obtained by only approximately 20% of the entire screened sample. Similarly, the mean standard score of CAL/WP-MD students was 85.88 on the Arithmetic subtest of the Wide Range Achievement Test 3 (WRAT; Wilkinson, 1993). Mean standard score on KeyMath–Revised (KM; Connolly, 1998) Problem Solving was 97.38, although this measure has a floor effect at the beginning of second grade, which makes it difficult for students to register low standard scores. Scores in the range noted previously indicate high risk for poor outcomes, based on a previous study with a similar sample (Fuchs, Zumeta, et al., 2010). For the 265 CAL/WP-MD students, 96 were in CAL-RTI (36%), 91 in WP-RTI (34%), and 78 in CON (29%). Their teachers were similarly distributed (34.1%, 36.5%, and 29.1%, respectively), reflecting the design of the parent study. See Table 1 for demographic information by tutoring condition. Students' mean age was 7.57 years (SD = 0.40). For a breakdown of demographic information by cohort, contact the first author.

Measures

Screening. We used three measures to determine high-risk status. We administered four subtests of the Second-Grade Calculations Battery (SGCB; Fuchs, Hamlett, & Powell, 2003): Sums to 12, Sums to 18, Minuends to 12, and Minuends to 18. All four subtests require single-digit addition or subtraction. Students have 1 min to complete up to 25 problems. Score is the number of correct answers. Alpha ranged from .85 to .93. We used Sums to 12 as the CAL screening measure for determining high-risk students. We used the other three subtests to index CAL outcomes (see next section). Story Problems (Jordan & Hanich, 2000) comprises 14 additive WPs. Problems represent combine, compare, and change WP types and require single-digit addition or subtraction for solution. The test is untimed. The tester reads each WP aloud as the students follow along on paper, and then the tester provides time for students to write answers and do any WP work. A student's score is the number of correct answers. Alpha was .87. The WASI (Psychological Corporation, 1999) is an individually administered measure of general cognitive ability with two subtests. Vocabulary measures word knowledge and expressive language with four picture items and 37 vocabulary words. Students define words until the end of the test or until reaching a ceiling of five consecutive errors. As reported by Psychological Corporation (1999), split-half reliability is .86. Matrix Reasoning measures nonverbal reasoning with pattern completion, classification, analogy, and serial reasoning tasks. Students choose the best of five choices to complete a visual patterns for 35 items or until reaching a ceiling of four consecutive errors or four errors over five consecutive items. Reliability is .94. The two subtest scores combine to yield an estimated full-scale IQ score with reliability of .92.

CALs. We measured proximal single-digit CAL outcomes with the three nonscreening subtests of the SGCB (Fuchs et al., 2003). The outcome was the sample-based average of these three measures. To measure proximal double-digit CALs, we administered Two-Digit Addition and Two-Digit Subtraction (Fuchs et al., 2003). For each two-digit test, students have 3 min to complete up to 20 two-digit plus and minus two-digit problems. Alphas for Two-Digit Addition and Two-Digit Subtraction were .96 and .87, respectively. The outcome used was the sample-based average of these two measures. We considered single-digit and double-digit CALs separately, given the overall low pretest skill of these at-risk students.

We measured distal effects with WRAT Arithmetic (Wilkinson, 1993), KM Addition, and KM Subtraction (Connolly, 1998). WRAT was administered in groups; KM, individually. On WRAT, students work on 40 CALs of increasing difficulty including addition, subtraction, multiplication, and division of whole numbers and fractions. Students who do not meet a basal on the written arithmetic also complete 15 oral arithmetic problems. Score is the number of correct answers. Alpha was .93. On KM Addition, students answer six problems with pictures and 12 problems in a written format. Problems range from one-digit addition to addition with decimals, fractions, and negative numbers. KM Subtraction functions similarly with six problems presented with graphics and 12 written problems of one-, two-, and three-digit subtraction with whole numbers, fractions, decimals, and negative numbers. Both KM subtests are scored by correct answers. Alphas for the two KM subtests were .84 and .81, respectively. Within this sample, analysis indicated a single-factor solution across the three tests (WRAT Arithmetic, KM Addition, and KM Subtraction) was a good fit to the posttest data, and therefore we used this factor score as the distal CALs outcome.

WPs. We used two outcomes to measure WP performance: an experimental measure of proximal effects and a factor score of commercial measures for measuring distal effects. The proximal measure and distal composite correlated .52. Second-Grade Story Problems (SGSP; Fuchs, Seethaler, et al., 2008), the proximal measure, includes 18 WPs (never used for instruction) representing combine, compare, and change problems types, with missing information in all three positions of the problem schema, with and without irrelevant information, with and without charts or graphs, and with and without multiple steps. Solutions require one-digit addition or subtraction. Five items involve combine problems, six involve compare problems, six involve change problems, and one item is a two-step problem involving both combine and compare. In groups, the tester reads each problem aloud as students follow along on paper. Students have 1 min to write a constructed response before the tester reads the next problem. Students earn credit for correct WP CAL (i.e., math answer) and label to reflect students' processing of the problem statement and understanding of the problem's focus. Alpha was .88.

We measured distal effects with KM Problem Solving (Connolly, 1998) and the Iowa Test of Basic Skills (ITBS; Hoover, Hieronymous, Dunbar, & Frisbie, 1993) Data Interpretation and Problem Solving. KM Problem Solving includes 18 WPs of increasing difficulty, which involve all four operations on WPs with transparent solution strategies, nonroutine problems without clear solution strategies, and items requiring students to demonstrate comprehension of a WP without solving it.

Administration is individual, and items are presented orally with accompanying pictures. Students answer with an oral response. Testing is discontinued after three consecutive errors. Score is the number of correct multiple-choice answers. Alpha was .74. ITBS was administered in groups. Students work on eight WPs presented orally and 22 written problems. Problems include one- and two-digit CALs with and without irrelevant information, with and without charts or graphs, and with and without multiple steps. The tester reads all problems aloud, and students respond using a multiple-choice format. Alpha was .81. These two measures (i.e., KM Problem Solving and ITBS) correlated $r = .43$ and were combined into a sample-based composite score.

Prealgebraic Knowledge. Prealgebraic knowledge outcomes included proximal and distal measures administered only at posttest. The proximal measures were Find X and Number Sentences. Each is administered in groups, requires constructed responses, and begins with the tester modeling a sample problem. With Find X (Fuchs, Powell, et al., 2009), students solve equations ($a + b = c$ or $d - e = f$) that vary the position of x across all three positions. The score is the number of correctly solved equations. Alpha on this sample was .91. With Number Sentences (Fuchs, Powell, et al., 2009), the tester reads eight WPs aloud; students have 30 s to write the mathematical equation representing the problem structure (students do not find solutions). The score is the number of correct equations. Alpha was .84. These measures (i.e., Find X and Number Sentences) correlated $r = .47$ and were combined into a composite score.

For Cohort 1, the distal measure was Dynamic Assessment of Algebraic Knowledge (DA; see Fuchs, Compton, et al., 2008, for details), an individually administered measure of students' responsiveness to instruction on finding the missing variable in addition expressions (Skill A; e.g., $x + 5 = 11$ or $6 + x = 10$), simple multiplication expressions (Skill B; e.g., $3x = 9$), and equations with two missing variables (Skill C; e.g., $x + 2 = y - 1$; $y = 9$). Mastery of each skill is assessed before instructional scaffolding begins and recurs after each level of scaffolding. If mastery occurs, the tester administers a generalization problem (for Skill A, $3 + 6 + x = 11$; for Skill B, $14 = 7x$; for Skill C, $3 + x = y + y$; $y = 2$) and moves to the next skill. If mastery does not occur, the tester provides the first (or next) level of instructional scaffolding, which is followed by the mastery test. Each level of scaffolding increases instructional explicitness and concreteness. If a student fails to achieve mastery after all five scaffolding levels for a given skill, testing is terminated. Scores range from 0 to 21 (0 indicates students never mastered any skill; 21 indicates students mastered each skill on the pretest and got each generalization problem correct). Alpha was .84. The outcome was a sample-based z score.

In Cohorts 2 through 4, a composite was generated across DA and the Test of Pre-Algebraic Reasoning (TPAR; Fuchs, Seethaler, & Powell, 2009), which comprises two types of problems. The first problem types (20 items) are mathematical equivalence statements with letters (i.e., x and y) representing missing quantities: 18 in nonstandard format (e.g., $y + 4 = 9$) and two in standard format (i.e., $1 + 5 = x$). The next problem types (six items) are function tables, each of which shows a two-column table. The first column shows a variable (x or y); the second shows a function involving that variable; each row shows a value for the variable and the resulting value for the function. In one row, the value of the function is empty and shaded; the task is to complete that row. In groups, the tester demonstrates how to complete a sample problem for each problem type. Students have 8 min to complete the first problem type with mathematical equations and as much time as needed (until all but two students finish) to complete the second problem type with function tables. The correlation between the two problem types was .54. The pattern of results was the same for the two problem types, so we used the total score. Alpha was .88. In Cohorts 2 through 4, the score was a sample-based z score composite of DA and TPAR.

Study Conditions

Students participated in two-tiered CAL intervention, two-tiered WP intervention, or business-as-usual control. Tier 1 comprised 34 whole-class lessons with two lessons per week over a 17-week period. Each lesson lasted approximately 40 to 45 min. Research assistants (RAs) delivered all lessons in the students' classrooms. Tier 2 comprised 39 small-group tutoring lessons. Tutoring started at the beginning of Weeks 4 or 5 of the Tier 1 intervention and lasted for 13 weeks with three lessons a week. RAs provided tutoring in groups of two to three students, and each lesson lasted 25 to 30 min. Tier 2 tutoring took place outside of the student's classroom (i.e., library, conference room, hallway), and Tier 2 tutoring did not occur at the same time in the school day as Tier 1 intervention.

Instructional Framework to Support Connections With Prealgebra

We relied on Pillay et al.'s (1998) model to design the CAL-RTI instructional whole-class and tutoring programs and to design the WP-RTI instructional whole-class and tutoring programs in ways that potentially support the transition from arithmetic to algebra. In this section, we provide an overview of the guiding frameworks. For information on the specifics of the programs, consult Fuchs et al. (in press). Contact lynn.a.davies@vanderbilt.edu for program manuals under the title *Math Wise* for CAL whole-class instruction and tutoring and *Pirate Math* for WP whole-class instruction and tutoring (*Math Wise* because the theme is owls, with allusions to becoming "wiser" in mathematics by using appropriate CAL strategies; *Pirate Math* because the theme is pirates, with allusions to finding *x*, the unknown, in WPs, just as *X* marks the treasure on pirate maps).

CAL Framework. CAL whole-class intervention and small-group tutoring incorporated two areas of emphasis for developing conceptual and procedural competence with one- and two-digit addition and subtraction CALs (e.g., Fuchs et al., 2013; Fuson & Kwon, 1992; Geary et al., 2008; Groen & Resnick, 1977; LeFevre & Morris, 1999; Siegler & Shrager, 1984). We focused on interconnected knowledge about number (e.g., cardinality, inverse relation between addition and subtraction, commutative property). Students used connecting cubes to explore how a target number (e.g., 7) can be partitioned in different ways. For example, 7 can be made of 3 and 4 (3 + 4 = 7), and 2 can be taken away from 7 for a difference of 5 (7 – 2 = 5). Students focused on part-part-whole knowledge related to number families (e.g., 2 + 5 = 7, 5 + 2 = 7, 7 – 5 = 2, 7 – 2 = 5) and grouped families using manipulatives or graphics on paper to demonstrate how and why problems constitute a number family. Students also used number families to explore the inverse relation between addition and subtraction. We also included a focus on place value concepts of 10s and 1s. Students practiced counting by 10s, learned about the meaning of 0 in the 10s and 1s places, and used and regrouped base-10 manipulatives to represent one- and two-digit numbers.

The other area for emphasis in CAL-RTI was practice. Students learned counting strategies for solving problems involving two digits or one digit plus one digit without regrouping (i.e., addition) and those involving two digits or one digit minus one digit without regrouping (i.e., subtraction). Practice required students to generate many correct responses to such problems to help them form long-term representations to support retrieval. Practice helped build fluency, which, in turn, helped improve double-digit addition and subtraction. RAs also explicitly taught efficient procedures for identifying when regrouping was required in addition and subtraction problems and how regrouping differed in addition and subtraction.

CAL intervention was divided into six units: (a) equal sign as a relational term, (b) addition concepts and operational strategies for problems for which retrieval is a viable strategy (problems where both operands are one digit or one operand is one digit and the other is two digits but regrouping is not required), (c) concepts and operational strategies for similar problems involving subtraction, (d) concepts and operational strategies for addition problems with regrouping, (e) concepts and operational strategies for subtraction problems with regrouping, and (f) review (although cumulative review was also integrated throughout the first five units).

WP Framework. WP intervention incorporated Kintsch and colleagues' framework in which WP solving is thought to be an interaction between problem solving and language comprehension (Cummins, Kintsch, Reusser, & Weimer, 1988; Kintsch & Greeno, 1985; Nathan, Kintsch, & Young, 1992). Their model of WP solving suggests general features of understanding text apply across all WPs but that WPs differ by problem type. Students can learn to apply a problem-type schema to assist with solving specific WPs, and at second grade, the three additive schemas include combine, compare, and change (Riley & Greeno, 1988).

At the beginning of instruction, teachers taught a general WP attack strategy to use on all WPs: read the problem, underline the labels, and name the problem type (i.e., schema; "RUN"). After teaching the attack strategy, we focused on teaching students how to recognize the underlying structure of a WP and to use one of three schemas to solve a WP. Initially, teachers introduced each schema by presenting a WP without missing information to enable students to focus on the WP structure. Teachers asked students to focus on the action or lack of action in the WP as well as the content of the WP question. We used graphics to demonstrate the relations underlying each schema, and we provided students with the opportunity to place known information into the graphics. After learning about the underlying structure of a WP, students were taught "meta-equations" that represented each schema and could be used to plug in known information from the WP. The meta-equation for combine problems was P1 + P2 = T (i.e., Part 1 plus Part 2 equals the total). The meta-equation for compare problems was B – s = D (i.e., the bigger amount minus the smaller amount equals the difference), and the change meta-equation was St ± C = E (i.e., the start amount has an increased or decreased change for a new end amount). After determining the WP schema, students wrote the meta-equation and filled in known WP information under specific parts of the meta-equation. Teachers also taught students to represent missing WP information with an x.

We also provided explicit instruction on the language comprehension necessary for combine, compare, and change problem types. As in Jitendra (2007), we focused on helping students develop understanding of relational terms for compare WPs (e.g., *more–less, older–younger, hotter–colder*). Also to extend Jitendra (2007), we also focused on relational terms in static (i.e., combine and compare) versus change WPs (e.g., "Claire had 7 fewer candies than Winston" versus "Claire had 7 candies and gave 2 to Winston") and vocabulary related to quantities (e.g., *amount* refers to a quantity) and superordinate levels (e.g., "4 cows and 9 horses equals 13 animals") for combine WPs. Note, however, that we did not teach students "key words," which consensus suggests is not a helpful WP solving strategy (Clement & Bernhard, 2005). Instead, we focused on the deeper understanding of specific words and phrases for WP interpretation.

WP intervention was divided into five units: (a) foundational skills for the WP content (i.e., equal sign as a relational term, strategies to find x, strategies for checking WP work), (b) combine problems, (c) compare problems, (d) change problems, and (f) review (although cumulative review was also integrated throughout the first four units). The program typically provides explicit conceptual and strategy instruction on one- and two-digit CALs (e.g., Fuchs et al., 2009), but for the present study, we removed all instruction on CALs. When students asked questions or needed corrective feedback on CALs, they were told to use the strategies they learned from their classroom teachers.

Linkages With Prealgebraic Knowledge. Both CAL and WP intervention linked to prealgebraic knowledge, as per Pillay et al. (1998), in two ways. First, in both CAL and

WP intervention, we focused on equal-sign understanding. Specifically, teachers taught students to understand the equal sign as a relational symbol, indicating the two expressions on either side of the equal sign balanced (Jacobs, Franke, Carpenter, Levi, & Battey, 2007). Although using terminology (e.g., *same as*) to describe the equal sign improved student understanding of the equal sign (Baroody & Ginsburg, 1983; Blanton & Kaput, 2005), explicit instruction on the location or meaning of the equal sign also improves student understanding (McNeil & Alibali, 2005; Rittle-Johnson & Alibali, 1999). As part of WP intervention, Powell and Fuchs (2010) determined that third-grade students with MD who had equal-sign instruction embedded within WP tutoring performed better than students without embedded equal-sign instruction on solving equations and some types of WPs.

Second, we used meta-equations to help organize WP work by schema (i.e., P1 + P2 = T; B – s = D; St ± C = E). Students learned to identify the WP type and write the corresponding meta-equation. Then, using the meta-equation as a guide, students filled in the appropriate numbers from the WP and placed an *x* in the equation to represent the missing information (i.e., WP answer). Teachers taught specific strategies for solving for *x*, which was shown to increase the prealgebraic reasoning of second-grade students (Fuchs, Zumeta, et al., 2010). As the WP intervention provided this additional prealgebraic component (i.e., using meta-equations), we expected WP students to demonstrate stronger prealgebraic reasoning than CAL students.

Teachers' Instruction. The primary mathematics curriculum of the school district was Houghton Mifflin Math (Greenes et al., 2005). All teachers used this curriculum to guide mathematics instruction in second grade. The content of the curriculum matched the CAL intervention in terms of teaching one- and two-digit addition and subtraction. The content also matched the WP intervention by focusing on combine, compare, and change WP types. In this way, control students received CAL and WP instruction similar to that of the two interventions. See Fuchs et al. (in press) for a discussion of the similarities and differences among CAL intervention, WP intervention, and the district math curriculum.

Structure of CAL and WP Whole-Class Intervention. Both CAL and WP whole-class instruction occurred for 17 weeks with two 40- to 45-min lessons each week. In both Math Wise and Pirate Math, RAs provided explicit instruction. Each lesson began with an explanation of worked examples. Then, the RAs used guided practice to demonstrate strategies on partially worked and unworked problems. Next, students worked in pairs as the RA circulated and provided feedback as necessary. Finally, students participated in an independent practice activity related to the lesson's skill. Cumulative review was systematically incorporated throughout both CAL and WP interventions.

In CAL whole-class instruction, this general framework is captured in four activities per lesson. The first, the Daily Lesson, focuses on six to eight problems on the day's topic. This involves the teacher presenting worked examples and moving to partially worked and then unworked problems, with constant student participation. The second activity is Partner Work, in which paired students (a higher and a lower performer) work in a structured manner on 12 problems on that day's topic, as coach and player. As validated with peer-assisted learning strategies (Fuchs et al., 1997), the higher-performing student begins as coach, who asks the player step-by-step questions to model the teacher's solution strategy and provides corrective feedback. This occurs on three problems. Then the player "thinks aloud" the solution strategy on the next three problems while the coach monitors and provides corrective feedback. Students then switch roles, so the lower performer asks his or her partner to model the solution strategy on Problems 7 through 9 while providing corrective feedback, and finally, the higher performer talks aloud the final three problems. The teacher circulates while providing assistance. The third activity is Time Owl (beginning Lesson 15), which presents a page of different types of CAL problems. The teacher provides a directive (e.g., "Solve addition problems that require regrouping"); then students have 1 min to find and complete that type of problem. At the end of 1 min, students switch Time Owls with partners and check answers. The last activity, Individual Practice, provides students with a practice sheet of 25 problems on the day's topic and cumulative review. At the end of 5 min, students switch papers to check answers. During the final three activities, students earn points for working with partners cooperatively and producing accurate work. The pair with the most points collects Math Wise folders.

WP whole-class instruction also includes four activities per lesson. The first is the Teacher-Led Problem, in which the teacher reviews and leads discussion about a problem from the previous day's lesson. Each student has a copy of that problem, which the teacher and students work through together. The second activity is the Daily Lesson, focused on that day's topic. This involves the teacher presenting worked examples and moving to partially worked and then to unworked problems, with constant student participation. The third activity is Partner Work, in which students work in pairs to solve two WPs on the day's topic. After ~8 min, the teacher shares answers with the class, against which the pair compares its work. The final activity is Individual Practice, in which students complete five find-*x* problems and one WP. As each student finishes, the teacher checks his or her work. Students earn points for each part of a correctly applied solution strategy and record scores on a "treasure map."

Structure of CAL and WP Tutoring. An RA, different from the whole-class RA, conducted all tutoring lessons. Tutoring did not occur during the same time slot as whole-class instruction to ensure that students participated in both tiers of intervention. Tutoring content was similar to whole-class content, except that tutoring targeted the most difficult concepts from whole-class instruction, used manipulatives more frequently, employed additional scaffolding from the RA, taught additional strategies, and incorporated self-regulation with tangible reinforcements. Throughout each tutoring lesson, RAs asked questions to individual students or the small group to check for engagement and understanding. RAs set a timer at random intervals (as outlined in each lesson's guide) and awarded a check if all students in the group were following directions and working hard when the timer beeped. At the end of each lesson, students earned individual checks for correctly answering "bonus problems." Students did not know which problems worked during tutoring would end up as bonus problems until the end of each lesson. This encouraged students to do their best mathematics work throughout the lesson. Students colored an owl (for Math Wise) or a pirate-themed picture (for Pirate Math) for each checkmark earned during the lesson. After a student colored 16 pictures, which typically took three to four tutoring lessons to complete, the student selected a small prize out of a treasure box.

Six activities occurred during each CAL tutoring lesson. Fist, students participated in Number Combination Flash Cards. For Lessons 1 through 13, students worked on addition flash cards. Starting with Lesson 14, subtraction flash cards were mixed in with the addition cards. For 1 min, the RA showed cards to the students in a round-robin format. If a student answered incorrectly, the RA encouraged the student to use a counting strategy to answer the number combination correctly. At the end of 1 min, the RA and students counted the correctly answered cards. Then, the RA shuffled the cards, and the students had another 1 min to try to beat their score. The group then graphed their highest flash card score on a graph. The second activity each lesson was the Number Combinations Lesson. The RA reviewed number knowledge content (e.g., cardinality, mental representation of the number line, relationship between addition and subtraction, commutative property of addition) and efficient counting strategies (i.e., min strategy for addition; missing addend strategy for subtraction). Lessons began with a focus on number combinations of +1 and –1, moving to +0 and –0, and then doubles. After doubles, lessons focused on number combinations in sets (e.g., all combinations with a sum of 5 or 5 as the minuend). The RA encouraged students to use manipulatives, such as snap cubes, to understand sets of combinations. The third activity each lesson was Owl Belly, an activity named by the students. Each student had a worksheet with eight owls, and each owl had a number on its belly (e.g., 11). The tutor read a number combination aloud (e.g., "four plus seven"), and the students had to color the owl with the correct answer. RAs presented only six problems, and students had to write one addition and one subtraction number combination for the two uncolored owl bellies. The fourth activity each lesson was Double-Digit Flash Cards. Similar to the first flash card activity, students responded in a round-robin format for 1 min and answered whether the problem on the flash card required addition or subtraction (Lessons 1 to 21) or whether the problem required regrouping and why (Lessons 22 to 39). The fifth activity was Double-Digit Lesson. Here, the RA reviewed stories and rhymes taught in CAL whole-class intervention, which taught place value and regrouping concepts. Students completed six double-digit problems with Base-10 blocks. The sixth activity each lesson was Paper-and-Pencil Review. Students had 1 min to independently solve 10 one-digit problems and 2 min to independently solve five two-digit problems. RAs provided corrective feedback as necessary.

Four activities occurred each WP lesson with the first activity, Game, differing by lesson topic. Students played the Find *X* Game (solving for *x* in equations), the Total Game (filling in the total meta-equation for orally presented problems), or the Difference Game (underlining and drawing inequality symbols between two compared amounts in a WP). During the Daily Lesson, RAs provided scaffolded schema instruction as students completed three to four WPs. Most WPs aligned with the unit's focus, but previously learned WP types were incorporated to help students practice discrimination among problem types and to provide cumulative review. With the third activity, Sorting Game, the RA read WPs and students responded, round-robin style, for 2 min as to the problem type of the WP. At the end of 2 min, the RA provided corrective feedback for incorrectly identified WPs. The fourth activity each lesson was Paper-and-Pencil Review. Students worked individually for 2 min on solving nine find-*x* problems and for 2 min on solving one WP. RAs provided feedback as necessary.

RAs. Across the four cohorts, 39 RAs provided tutoring to 143 groups of students in groups of two or three students, with several "groups" of one created after students left the tutoring project. A typical RA was 1 or 2 years beyond undergraduate education and studying for a graduate degree in an education-related field. To control for quality of RAs, every RA taught or tutored in both CAL and WP intervention. We guarded against contagion of conditions by coding lesson guides and materials by condition, reviewing the differences between the two conditions in weekly meetings, and providing corrective feedback during in-school observations and digital recordings of small-group tutoring lessons.

To prepare RAs for tutoring duties, we conducted an introductory meeting during which we discussed the ethics of conducting school research. Then, during two 6-hr sessions, RAs learned how to implement CAL and WP whole-class teaching or small-group tutoring. We introduced RAs to the goals of the project, and we provided instruction and demonstration on implementing the lessons. RAs were paired together to practice delivery of lessons, and before providing tutoring in schools, RAs had to demonstrate proficiency with a project coordinator by delivering a lesson. RAs who achieved 95% fidelity against the lesson's fidelity checklist were judged ready for school tutoring. RAs who scored below 95% received coaching from the project coordinator, practiced lesson delivery, and then were reassessed at a later time on another lesson. During weekly meetings held throughout the year, the first author, project coordinators and

RA-tutors discussed upcoming tutoring lessons and any tutoring difficulties.

Fidelity. To assess fidelity, all sessions were digitally audio recorded, and we conducted in-school observations of each RA. Prior to the first tutoring session, we created checklists of essential information for each lesson within CAL intervention and WP intervention. At the end of each school year, RAs independently listened to a random sample of ~20% of tutoring tapes, representing conditions, RAs, tutoring groups, and lesson types equitably, while completing the checklists to identify the percentage of points addressed. For tutoring, the mean percentage of points addressed was 96.06 (SD = 2.83) for CAL and 96.34 (SD = 3.28) for WP, $t(35) = 0.58$, p = .563. (In these fidelity analyses, RA was the unit of analysis; t tests are for dependent samples because each RA taught in both CAL and WP conditions.)

Procedure

Over the four cohorts, we administered the screening measures in September and the pretest measures in October. RAs delivered whole-class instruction from November to March and tutoring from December to March. We posttested all students in late March. RAs independently entered responses on 100% of the test protocols for every screening and outcome measure on an item-by-item basis into two separate databases. We compared and rectified the discrepancies between the two databases to reflect the student's original response.

Analysis Plan

Multilevel modeling was utilized for primary analyses within SAS PROC MIXED, with restricted maximum likelihood estimation and reporting standard model fit indices (deviance or –2 log likelihood; Bayesian information criteria). We followed a model-building approach, where we first established the clustering effects, or intraclass correlation (ICC), of both teacher and school in unconditional models (Luke, 2002; Sullivan, Dukes, & Losina, 1999). Where the random effect of school was not contributory, we eliminated this term from the random portion of the model. Next, student-level fixed effects were added to the model, primarily the pretest variable associated with the dependent variable being analyzed. For CAL outcomes, the pretest covariates corresponded to the outcomes: single-digit SGCB, double-digit SGCB, and WRAT (for the standardized outcome). For WP outcomes, the pretest factor score was computed from SGSP and KM Problem Solving. Covariates for algebra outcomes were based on the strength of their relations with pretest variables: Proximal algebra was used for the WP factor score, distal algebra for the WP factor score and WRAT performance. Additional student-level covariates included gender, race or ethnicity, reduced-price or free lunch (RFL) status, special education status, and English-as-a-second-language (ESL) status, where these were related to outcomes. Treatment condition was then added as a teacher-level effect; cohort was also considered a teacher-level effect. Interactions among these terms, and interaction terms across tiers, were also considered and are discussed later where relevant.

To represent the size of the effect of treatment in addition to statistical significance, we used two measures. The primary standardized effect size (ES) was computed using parameter estimate differences between treatment groups, divided by the square root of the student-level residual variance, when treatment was the sole factor in the model. We also computed a proportional reduction in variance (PRV; Peugh, 2010; Singer & Willett, 2003) that is based on how well the final model does in reducing between-teacher variance, relative to a model with all other predictors except treatment; such values do not represent the absolute amount of variance explained in the outcome (Ma & Ma, 2008).

Results

Table 1 shows demographic information by tutoring condition. There were no significant differences on race or ethnicity, RFL status, special education status, ESL status, or retained status. There was a significant difference for gender, $\chi^2(2, N = 265) = 6.21$, $p = .045$. Table 2 shows means and standard deviations for each screening and outcome measure. For a breakdown of screening and outcomes measures by cohort, contact the author. There was no significant difference as a function of treatment condition on any pretest measure. Table 3 shows results for the two proximal CAL outcomes, Table 4 for the distal CAL outcome, Table 5 for the proximal and distal WP outcomes, and Table 6 for the proximal and distal prealgebraic knowledge outcomes. The top portion of each table shows results of the unconditional models, along with model fits. The bottom shows results for the final model, including model fit, effects for treatment, cohort, pretest, and demographically relevant covariates.

Proximal CAL Effects

For single-digit outcomes, the unconditional model suggested a random effect only for teacher, with a strong ICC of .44. The final model included teacher as a random effect; for fixed effects, at Level 1 (L1), once pretest was in the model, no other predictors at this level were significant. Level 2 (L2) teacher-level predictors were treatment and cohort; there was no interaction between pretest and treatment. In the final model, there were significant main effects for pretest ($p < .001$, with higher outcomes associated with higher pretest performance) and treatment condition, $F(2, 89.1) = 30.56$, $p < .001$. Cohort did not interact with treatment and was not a significant effect alone, $F(3, 148) = 2.36$, $p = .074$. Follow-up to the treatment effect indicated CAL-RTI outperformed WP-RTI ($p < .001$; ES = 0.95) and CON ($p < .001$; ES = 1.67), and WP-RTI outperformed CON ($p < .001$; ES = 0.71). Adding treatment condition resulted in a PRV of 61% across teachers, relative to the prior model with all predictors except treatment.

For double-digit CALs, the unconditional model also included only teacher as a random effect, with an ICC similar to single-digit CALs (.45). The final model included both teacher and school as random effects; fixed effects included pretest (at L1) and cohort and treatment (at L2). In the final model, there were no significant interactions. There were positive effects for pretest ($p = .003$); for cohort, $F(3, 50.6) = 3.37$, $p = .026$; and for treatment condition, $F(2, 64.9) = 35.65$, $p < .001$. CAL-RTI outperformed WP-RTI ($p < .001$; ES = 1.21) and CON ($p < .001$; ES = 1.67), but WP-RTI was comparable to control ($p = .089$; ES = 0.45). For double-digit CALs, PRV for treatment condition was 73% across teachers and schools. Thus, results were directionally similar to those found for single-digit CALs, except WP intervention was comparable to control on double-digit problems.

Distal CAL Effects

On the distal CAL outcome, the unconditional model revealed an ICC of .03 for school and .20 for teacher. The final model included both teacher and school as random effects; fixed effects were pretest, age, and special education status (at L1) and cohort and treatment (at L2). In the final model, there were no significant interactions. There were positive effects for pretest ($p < .001$), age ($p < .001$), special education ($p = .008$), and treatment condition, $F(2, 66.6) = 34.68$, $p < .001$, but not cohort ($p = .713$). CAL-RTI outperformed WP-RTI ($p < .001$; ES = 0.72) and CON ($p < .001$; ES = 1.19); WP-RTI also outperformed CON ($p = .005$; ES = 0.48). Adding treatment resulted in a PRV of 73% across teachers and schools for distal CALs.

Proximal WP Effects

For the unconditional model, the ICC for teacher was .43; school was noncontributory. The final model included both teacher and school as random effects; fixed effects were pretest, age, special education status (at L1) and cohort, treatment, and the interaction between cohort and treatment (at L2). In the final model, there was a significant interaction of cohort and treatment, $F(6, 73.2) = 5.77$, $p < .001$, that suggested smaller differences

Table 2. Performance Data by Treatment Condition.

	Condition											
	CAL				WP				Control			
	Pre		Post		Pre		Post		Pre		Post	
Variable	*M*	(*SD*)	*M*	(*SD*)	*M*	(*SD*)	*M*	(*SD*)	*M*	(*SD*)	*M*	(*SD*)
Screening												
Sums to 12	3.82	(1.64)	NA		4.04	(1.61)	NA		4.15	(1.53)	NA	
Story Problems	2.89	(1.60)	NA		3.15	(1.51)	NA		3.12	(1.51)	NA	
WASI IQ	85.54	(7.29)	NA		87.55	(9.01)	NA		86.65	(8.72)	NA	
CAL outcomes												
Proximal												
Sums to 18	4.15	(2.93)	12.09	(4.53)	3.70	(2.66)	8.21	(4.37)	4.26	(2.97)	6.19	(3.71)
Minuends to 12	2.98	(2.18)	8.53	(4.64)	3.04	(1.90)	6.86	(3.21)	2.95	(2.02)	4.65	(3.12)
Minuends to 18	2.28	(2.31)	8.07	(3.82)	2.07	(2.03)	4.87	(3.17)	2.05	(2.12)	3.13	(2.39)
Two-Digit Addition	1.90	(2.45)	14.92	(5.10)	1.57	(2.09)	8.50	(4.91)	1.56	(1.72)	6.72	(5.27)
Two-Digit Subtraction	0.98	(1.64)	7.16	(4.80)	1.04	(1.57)	4.79	(3.12)	0.86	(1.71)	3.65	(3.07)
Distal												
WRAT Arithmetic[a]	87.28	(12.61)	101.39	(9.58)	85.42	(11.20)	95.54	(9.96)	84.67	(11.91)	90.71	(10.94)
KM Addition[a]	NA		106.63	(12.41)	NA		98.02	(10.69)	NA		94.10	(12.45)
KM Subtraction[a]	NA		97.37	(15.07)	NA		91.98	(12.42)	NA		89.68	(14.33)
WP outcomes												
Proximal												
SGSP	4.92	(2.88)	7.91	(3.89)	5.07	(2.24)	16.05	(8.11)	4.30	(2.18)	8.59	(4.39)
Distal												
KM Problem Solving[a]	97.81	(8.73)	98.62	(9.68)	97.20	(8.00)	99.84	(8.38)	97.05	(8.54)	98.53	(9.23)
ITBS[b]	NA		25.75	(13.58)	NA		24.69	(12.74)	NA		24.10	(12.14)
Prealgebraic reasoning												
Proximal												
Find X	NA		3.39	(2.90)	NA		6.20	(1.93)	NA		3.13	(2.83)
Number Sentences	NA		0.95	(1.22)	NA		3.11	(2.27)	NA		0.69	(1.14)
Distal												
DA	NA		5.43	(3.83)	NA		5.87	(3.99)	NA		4.44	(3.84)
TPAR	NA		8.28	(3.66)	NA		9.77	(3.81)	NA		7.75	(4.03)

Note. CAL = calculation intervention; WP = word problem intervention; WASI = Wechsler Abbreviated Scale of Intelligence (Psychological Corporation, 1999); WRAT = Wide Range Achievement Test (Wilkinson, 1993); KM = KeyMath–Revised (Connelly, 1998); SGSP = Second-Grade Story Problems (Fuchs, Seethaler, et al., 2008); ITBS = Iowa Test of Basic Skills (Hoover, Hieronymous, Dunbar, & Frisbie, 1993); DA = Dynamic Assessment of Algebraic Knowledge (Fuchs, Compton, et al., 2008); TPAR = Test of Pre-Algebraic Reasoning (Fuchs, Seethaler, & Powell, 2009).
[a]Standardized score.
[b]Normal curve equivalent.

between the WP condition and the other conditions in the first two cohorts relative to the third and fourth cohorts. Even in the context of this interaction, there were positive effects for pretest ($p < .001$), special education ($p = .038$), cohort ($p = .002$), and treatment condition,

Table 3. Model Characteristics for Proximal Calculation Outcomes.

	Single-digit proximal					Double-digit proximal				
Parameter	Estimate	(*SE*)	*df*	*t*/*z*/*F*	95% CI	Parameter estimate	(*SE*)	*df*	*t*/*z*/*F*	95% CI
Unconditional model										
Fixed: Intercept (γ_{000})	6.97	(0.31)	103	22.32***	[6.35, 7.59]	7.67	(0.40)	93.3	19.25	[6.88, 8.47]
Random: Teacher (σ_{u0}^2)	5.91	(1.27)		4.65***	[4.04, 9.47]	9.79	(2.22)		4.41***	[6.56, 16.25]
School (σ_{v0}^2)			NA					NA		
Residual (σ_e^2)	7.63	(0.80)		9.57***	[6.28, 9.47]	11.98	(1.30)		9.25***	[9.80, 14.99]
Model fit	Dev = 1392.2; Param = 2; BIC = 1401.4					Dev = 1503.7; Param = 2; BIC = 1512.9				
Full/final model										
Fixed: Intercept (γ_{000})	2.87	(0.63)	134	4.58***	[1.63, 4.11]	3.14	(0.93)	110	3.36**	[1.29, 4.99]
Pretest	0.53	(0.10)	234	5.25***	[0.33, 0.73]	1.69	(0.57)	246	2.98**	[0.57, 2.80]
Treatment (CAL)	4.45	(0.58)	85.1	7.69***	[3.31, 5.60]	5.57	(0.71)	64	7.81***	[4.15, 7.00]
Treatment (WP)	1.86	(0.58)	91	3.17[b]	[0.70, 3.02]	1.24	(0.72)	66.4	1.72	[–0.20, 2.68]
Cohort 1	–0.38	(0.62)	110	–0.62	[–1.62, 0.85]	–0.68	(0.80)	35.5	–0.85	[–2.31, 0.95]
Cohort 2	0.73	(0.65)	136	1.13	[–0.55, 2.02]	0.39	(0.87)	42.8	0.45	[–1.36, 2.13]
Cohort 3	1.09	(0.62)	133	1.76	[–0.14, 2.31]	1.63	(0.79)	26.6	2.06*	[0.01, 3.26]
Random: Teacher (σ_{u0}^2)	2.01	(0.70)		2.88**	[1.12, 4.57]	2.07	(1.32)		1.57	[0.80, 12.70]
School (σ_{v0}^2)			NA			0.16	(0.63)		0.25	[0.01, 5.80]
Residual (σ_e^2)	6.84	(0.72)		9.50***	[5.62, 8.50]	12.10	(1.33)		9.11***	[0.44, 0.53]
Model fit	Dev = 1279.4; Param = 2; BIC = 1288.6					Dev = 1378.2; Param = 3; BIC = 1387.9				

Note. Pretest refers to single-digit and double-digit computation, respectively. For the unconditional model, *t* values are reported. Other variables considered but not needed in final model (sex, lunch status, ethnicity, second-language status, special education status, age) are not shown. For full/final model, *F* values are reported for the fixed portion of model, *z* values for the random portion of model. For *df*s, numerators (not included) are all 1; denominators (reported) use the Kenward-Roger approximation. Under fixed effects, intercept refers to the outcome variable. Pretest is grand-mean centered. The code 0 is for the treatment control group and for Cohort 4. Dev = deviance or –2 log likelihood;
Param = (random) parameters; BIC = Bayesian information criterion; CAL = calculation intervention; WP = word problem intervention. For Dev. and BIC, lower values are better fit. Intraclass correlations are reported in text.
*$p < .05$. **$p < .01$. ***$p < .001$.

Table 4. Model Characteristics for Distal Calculation Outcome.

	Distal (standardized) calculation				
Parameter	Parameter estimate	(*SE*)	*df*	*t*/*z*/*F*	95% CI
Unconditional model					
Fixed: Intercept (γ_{000})	–0.00	(0.07)	21.9	–0.01	[–0.14, 0.15]
Random: Teacher (σ_{u0}^2)	0.14	(0.06)		2.38[b]	[0.07, 0.38]
School (σ_{v0}^2)	0.02	(0.03)		0.71	[0.00, 24.13]
Residual (σ_e^2)	0.53	(0.06)		9.55[c]	[0.43, 0.65]
Full/final model					
Model fit		Dev = 636.2; Param = 3; BIC = 645.9			
Fixed: Intercept (γ_{000})	0.57	(0.92)	244	0.63	[–1.23, 2.38]
Pretest	0.03	(0.00)	245	8.66***	[0.02, 0.04]
Age	–0.44	(0.10)	243	–4.47***	[–0.63, –0.25]
Special education (receive)	–0.31	(0.11)	244	–2.67**	[–0.54, –0.08]
Treatment (CAL)	0.73	(0.09)	67.8	8.07***	[0.55, 0.91]
Treatment (WP)	0.26	(0.09)	74.9	2.90**	[0.08, 0.44]
Cohort 1	–0.11	(0.13)	36.2	–0.89	[–0.37, 0.15]
Cohort 2	–0.04	(0.14)	45.2	–0.27	[–0.31, 0.24]
Cohort 3	0.01	(0.13)	33.8	0.06	[–0.26, 0.28]
Random: Teacher (σ_{u0}^2)	0.00	(0.02)		0.31	[0.00, –0.00]
School (σ_{v0}^2)	0.03	(0.02)		1.64	[0.01, 0.15]
Residual (σ_e^2)	0.27	(0.03)		9.70***	[0.22, 0.33]
Model fit		Dev = 436.0; Param = 3; BIC = 445.6			

Note. See note for Table 3. The pretest was Wide Range Achievement Test Arithmetic (Wilkinson, 1993). The code 0 is for the treatment control group, for no special education, and for Cohort 4.
**p* < .05. ***p* < .01. ****p* < .001.

F(2, 40.7) = 44.77, *p* < .001. WP-RTI outperformed CAL-RTI (*p* < .001; ES = 1.47) and CON (*p* < .001; ES = 1.31), but CAL-RTI and CON were comparable (*p* = .152; ES = –0.16). In Cohorts 3 and 4, the pattern of treatment conditions fit the overall pattern. However, in Cohort 1, students in WP-RTI outperformed CAL-RTI but not CON; in Cohort 2, there were no significant differences. By contrast, in Cohorts 1 and 2, WP-RTI outperformed CAL-RTI and CON, but CAL-RTI and CON were comparable. The PRV for treatment was 88% across teachers and schools for proximal WPs.

Distal WP Effects

On distal WP outcomes, the unconditional model revealed an ICC of .06 for school; the ICC for teacher was noncontributory. For completeness, the remaining steps were followed and presented here, although given little variability in outcomes across teachers, treatment effects (assigned at the teacher level) are unlikely. The final model included school as a random effect; fixed effects were pretest and age (at L1) and cohort and treatment (at L2). In the final model, there were no significant interactions. There were positive effects for pretest (*p* < .001) and age (*p* < .001) but not for cohort (*p* = .304) or treatment (*p* = .871). The ES for WP-RTI versus CAL-RTI (0.05) and CON (ES = 0.15) were negligible, as was the ES for CAL-RTI versus CON (0.09). The PRV for the addition of treatment was <1%.

Proximal Prealgebraic Knowledge Effects

For the unconditional model, the ICC for teacher was .37; school was noncontributory. The final model included teacher as a random

Table 5. Model Characteristics for Proximal and Distal Word Problems Outcomes.

	Proximal					Distal				
Parameter	Estimate	(*SE*)	*df*	*t*/*z*/*F*	95% CI	Parameter estimate	(*SE*)	*df*	*t*/*z*/*F*	95% CI
Unconditional model										
Fixed: Intercept (γ_{000})	10.98	(0.59)	87.3	18.72***	[9.82, 12.15]	–0.01	(0.06)	24.2	–0.12	[–0.12, 0.11]
Random: Teacher (σ_{u0}^2)	20.54	(5.00)		4.11***	[13.42, 35.34]			NA		
School (σ_{v0}^2)			NA			0.03	(0.02)		1.51	[0.01, 0.19]
Residual (σ_e^2)	27.56	(3.01)		9.17***	[22.50, 34.55]	0.40	(0.04)		11.07***	[0.34, 0.48]
Model fit	Dev = 1729.0; Param = 2; BIC = 1738.1					Dev = 520.9; Param = 2; BIC = 527.3				
Full/final model										
Fixed: Intercept (γ_{000})	6.46	(1.28)	65.3	5.06***	[3.91, 9.01]	2.77	(0.69)	246	4.01***	[1.41, 4.13]
Pretest	0.73	(0.12)	236	6.12***	[0.50, 0.97]	0.53	(0.06)	245	8.36***	[0.41, 0.66]
Special education (receive)	–2.20	(1.05)	241	–2.09*	[–4.27, –0.13]			NA		
Age			NA			–0.38	(0.09)	246	–4.19***	[–0.55, –0.20]
Treatment (CAL)	–2.38	(1.55)	71.7	–1.54	[–5.47, 0.71]	–0.05	(0.09)	246	–0.50	[–0.22, 0.13]
Treatment (WP)	6.82	(1.72)	88.9	3.96***	[3.40, 10.23]	–0.01	(0.09)	246	–0.16	[–0.20, 0.17]
Cohort 1	–0.89	(1.71)	50.9	–0.52	[–4.32, 2.54]	0.11	(0.12)	38.6	0.86	[–0.14, 0.36]
Cohort 2	–1.61	(1.84)	61	–0.88	[–5.29, 2.06]	0.23	(0.13)	46.4	1.76	[–0.03, 0.50]
Cohort 3	–1.31	(1.83)	61.9	–0.72	[–4.98, 2.35]	0.06	(0.13)	35	0.51	[–0.19, 0.32]
Random: Teacher (σ_{u0}^2)	0.81	(2.02)		0.40	[0.10, 2.28]			NA		
School (σ_{v0}^2)	1.18	(1.19)		0.99	[0.32, 49.06]	0.02	(0.01)		1.43	[0.01, 0.16]
Residual (σ_e^2)	22.73	(2.57)		8.86***	[18.43, 28.7]	0.30	(0.03)		10.77***	[0.25, 0.36]
Model fit	Dev = 1516.3; Param = 3; BIC = 1526.0					Dev = 440.7; Param = 2; BIC = 447.2				

Note. See note for Table 3. Pretest refers to Second-Grade Story Problems (proximal) or Word Problems Standardized Composite (distal). Estimates for the significant treatment by cohort interaction in left-hand model not shown, but were included in model that produced the reported numbers; results described in text. The code 0 is for the treatment control group, for no special education, and for Cohort 4.

* $p < .05$. ** $p < .01$. *** $p < .001$.

Table 6. Model Characteristics for Proximal and Distal Pre-Algebra Outcomes.

Parameter	Proximal					Distal				
	Estimate	(*SE*)	*df*	*t*/*z*/*F*	95% CI	Estimate	(*SE*)	*df*	*t*/*z*/*F*	95% CI
Unconditional model										
Fixed: Intercept (γ_{000})	0.01	(0.06)	92.3	0.16	[–0.10, 0.12]	-0.04	(0.06)	20.6	–0.67	[–0.18, 0.09]
Random: Teacher (σ_{u0}^2)	0.17	(0.04)		3.88***	[0.11, 0.30]	0.09	(0.05)		1.61	[0.03, 0.49]
School (σ_{v0}^2)			NA			0.01	(0.02)		0.21	[0.00, 7.30]
Residual (σ_e^2)	0.29	(0.03)		9.30***	[0.24, 0.37]	0.67	(0.07)		9.71***	[0.55, 0.82]
Model fit	Dev = 511.3; Param = 2; BIC = 520.4					Dev = 673.9; Param = 3; BIC = 683.6				
Full/final model										
Fixed: Intercept (γ_{000})	–0.30	(0.09)	98.8	–3.40***	[–0.47, –0.12]	–1.10	(0.41)	239	–2.72**	[–1.90, –0.30]
Word Problem Pretest	0.24	(0.06)	247	3.99***	[0.12, 0.36]	0.48	(0.10)	247	5.04***	[0.29, 0.67]
WRAT 3 Pretest			NA			0.01	(0.00)	246	2.50*	[0.00, 0.02]
Special education (receive)	–0.31	(0.11)	247	–2.78**	[–0.53, –0.09]	–0.46	(0.16)	247	–2.81**	[–0.79, –0.14]
Treatment (CAL)	0.02	(0.09)	80.5	0.25	[–0.16, 0.20]	0.10	(0.13)	69.5	0.73	[–0.17, 0.36
Treatment (WP)	0.83	(0.09)	86.1	8.98***	[0.65, 1.02]	0.32	(0.13)	72.4	2.47*	[0.06, 0.59
Cohort 1	0.06	(0.10)	111	0.56	[–0.14, 0.26]	–0.23	(0.16)	44.2	–1.47	[–0.55, 0.09
Cohort 2	–0.06	(0.11)	123	–0.53	[–0.27, 0.16]	0.14	(0.17)	49.6	0.85	[–0.20, 0.48]
Cohort 3	0.08	(0.10)	126	0.79	[–0.12, 0.27]	0.05	(0.16)	37.6	0.32	[–0.27, 0.37]
Random: Teacher (σ_{u0}^2)	0.02	(0.02)		1.11	[0.01, 0.40]	0.03	(0.04)		0.79	[0.01, 8.30]
School (σ_{v0}^2)			NA			0.02	(0.02)		0.77	[0.00, 5.36]
Residual (σ_e^2)	0.26	(0.03)		9.61***	[0.21, 0.32]	0.53	(0.06)		9.47***	[0.43, 0.66]
Model fit	Dev = 412.6; Param = 2; BIC = 421.7					Dev = 600.8; Param = 3; BIC = 610.5				

Note. See note for Table 3. The code 0 is for the treatment control group, for no free or reduced lunch, for no special education, and for Cohort 4. WRAT = Wide Range Achievement Test (Wilkinson, 1993).
*p < .05. **p < .01. ***p < .001.

effect, and fixed effects were pretest and special education status (at L1) and cohort and treatment (at L2). There were positive effects for pretest ($p < .001$), special education ($p = .006$), and treatment condition, $F(2, 81.1) = 58.38$, $p < .001$, but not for cohort ($p = .538$). WP-RTI outperformed CAL-RTI ($p < .001$; ES = 1.52) and CON ($p < .001$; ES = 1.67), but CAL-RTI and CON were comparable ($p = .802$; ES = 0.16). The PRV for treatment condition was 88% across teachers.

Distal Prealgebraic Knowledge Effects

For the unconditional model, the ICC for teacher was .11; for school, the ICC was .01. The final model included teacher as a random effect, and fixed effects were pretests (both WP composite and WRAT 3 computations) and special education status (at L1) and cohort and treatment (at L2). There were positive effects for WP pretest ($p < .001$), WRAT 3 ($p < .013$), special education ($p = .005$), and treatment condition, $F(2, 68) = 3.38$, $p = .040$, but not for cohort ($p = .099$). WP-RTI outperformed control ($p = .016$; ES = 0.52). However, WP-RTI was comparable to CAL-RTI ($p = .063$), with an ES of 0.22. CAL-RTI and CON were also comparable ($p = .470$; ES = 0.25). Adding treatment condition results in a PRV of 19% across teachers and schools relative to the prior model.

Discussion

In the present study, we focused on the efficacy of multitier mathematics RTI systems in CALs and WPs for second-grade students at risk for CAL-WP/MD. We compared students at risk for CAL-WP/MD who received two tiers of support (i.e., whole-class instruction and small-group tutoring) in either CALs or WPs to each other and to a group of business-as-usual control students who were also at risk for CAL-WP/MD. This was our major contrast, but we also included the business-as-usual condition to control for history and maturation effects and to determine whether students in the RTI conditions made more progress than would have occurred without RTI.

On both proximal and distal CAL outcomes, CAL-RTI students outperformed WP-RTI and CON students with ESs ranging from 0.72 to 1.67. Our CAL interventions, focused on establishing conceptual understanding and building fluency with single- and double-digit addition and subtraction CALs, proved effective. Interestingly, on the distal CAL outcome, WP-RTI students outperformed CON students (ES = 0.48), even as CAL-RTI students demonstrated superior performance over WP-RTI students (ES = 0.72). This result is not surprising as WP-RTI students performed CALs on WPs during every WP lesson. Our hypothesis related to CAL outcomes, that CAL-RTI would promote superior CAL outcomes, was corroborated. Because students with MD often demonstrate a lack of fluency with number combinations (Baroody, Bajwa, & Eiland, 2009), and given that difficulty with number combinations often contributes to difficulty with solving more complex computation problems (Chong & Siegel, 2008) and higher-level mathematics problems in the areas of rational numbers, algebra, and problem solving (Wei, Lenz, & Blackorby, 2012), this finding is also important: The CAL-RTI approach shows promise for establishing strong CAL skill for students with CAL/WP-MD.

On proximal WP measures, WP-RTI students demonstrated significant growth over CAL-RTI and CON students with ESs of 1.31 to 1.47. Interestingly, WP results from Cohorts 1 and 2 were not as strong as results from Cohorts 3 and 4. Between Cohorts 2 and 3, we designed a language component for both whole-class and small-group WP instruction, and this component was embedded within WP instruction for Cohorts 3 and 4. At the whole-class tier, RAs presented students with "compare sentences," and the RA and students worked together to identify a "compare word" to designate which item or person has the bigger or smaller amount. To reduce demands on working memory, students were taught to place an inequality symbol (i.e., $<$ or $>$) above the compare word. The RA then worked with the class to write a new compare sentence using an opposite compare word, for which the relationship between the bigger and smaller amounts was preserved. Within small-group tutoring, the language component included working on two compare sentences, placing the inequality symbol over the compare word, and writing *B* (for "bigger amount") and *s* (for "smaller amount") over the items or persons compared. Students then filled in the difference meta-equation (i.e., B – s = D) and placed *x* in place of the missing information. Although difficult to pinpoint the exact cause for better WP performance from Cohorts 3 and 4, the added focus on the language used in WPs is a likely contributor to better WP understanding by problem type or increased understanding of the relationship between numbers in WPs. The language component may be especially important for students with CAL/WP-MD because these students often experience comorbid language and reading difficulties (Fuchs & Fuchs, 2002). Even so, on distal WP measures, there were no significant differences among CAL/WP-MD students in the three conditions, and this result mirrors previous WP studies using commercial WP measures as outcomes (Fuchs, Zumeta, et al., 2010). And it is important to note that the proximal measure shared important design elements with commercial WP tests: It mixed different problem types while requiring transfer to untaught program components. Therefore, with strong WP effects on this proximal measure, we corroborate our hypothesis that WP-RTI positively increases WP outcomes. Even though successful WP performance may not necessarily translate to success with real-world problem solving, WPs are important because they represent a major emphasis in every strand of the mathematics curriculum (i.e., solving WPs using fractions, geometry, measurement, algebra, etc.) and because WP difficulty has implications for both successful school and occupational success (Murnane, Willett, Braatz, & Duhaldeborde, 2001; Parsons & Bynner, 1997). With a WP-RTI approach, the WP skill of students with CAL/WP-MD can be improved.

We were also interested in how CAL and WP intervention transferred to performance on prealgebraic measures. Confirming our hypothesis that WP-RTI would provide a stronger route than CAL-RTI to prealgebraic reasoning, WP-RTI students outperformed CAL-RTI and CON students on proximal measures of prealgebraic reasoning with ESs of 1.52 to 1.67. Moreover, CAL-RTI and CON students performed similarly. On distal measures of prealgebraic reasoning, WP-RTI demonstrated a significant gain over CON students (ES = 0.52) and a marginally significant gain over CAL-RTI students (ES = 0.22). Similar to proximal measures of prealgebraic reasoning, CAL-RTI and CON students performed comparably. Our results therefore indicate WP-RTI is more influential than CAL-RTI in improving the prealgebraic reasoning of students with CAL/WP-MD. This finding adds to Fuchs et al.'s (2012) correlational study, in which the connection to algebra for WPs was stronger than CAL.

Our results indicate WP-RTI is more influential than CAL-RTI in improving the prealgebraic reasoning of students with CAL/WP-MD.

Kieran's (1990) developmental progression on algebraic competence suggests arithmetic competence is a foundational skill necessary for algebra performance. Kieran states that part of arithmetic competence involves understanding symbols and solving equations. Even though both CAL and WP interventions focused on using operator symbols (e.g., +, –, =) and solving equations, we have two potential explanations for the superior prealgebraic performance of students in the WP-intervention. First, the WP interventions required students to develop conceptual understanding attached to symbols presented through real-life scenarios (i.e., WPs). For example, students interpret "Annabelle lost eight flowers" as a subtraction expression, and "Lincoln bought some more candy" as an addition expression. By learning problems types by schema, students worked on understanding the underlying structure of an addition or subtraction number sentence. Second, the WP interventions taught students to set up and solve equations where a sum or difference was not always the required computation. For example, CAL students always solved equations such as $14 - 9 = \_$ or $3 + 8 = \_$. WP-RTI students solved equations such as $14 - 9 = x$ but also solved equations such as $3 + x = 11$ or $x - 4 = 7$. To solve these latter equations, students had to evaluate the missing part (i.e., x) and determine how to solve for x. Solving for x sometimes required not using the operator symbol in the equation (e.g., $x - 4 = 7$ is solved by adding 4 and 7 together). By solving equations with a missing variable, WP-RTI students worked on Kieran's second stage of algebraic competence where the focus is understanding the equal sign as relational and understanding the concept of a variable. These results have important implications for designing early intervention for students at risk for long-term MD, including difficulty with algebra. Findings indicate that an instructional focus on WPs, which is deliberately designed to forge connections with prealgebraic thinking (as in the present study), assists students in their transition to formal algebraic thinking. Finding this link with a sample of students with concurrent difficulty with CALs and WPs is notable.

Conclusion

In sum, we learned that a multitier support system focused on CALs improves the CAL outcomes of students at risk for CAL/WP-MD and that, to realize benefits on double-digit addition and subtraction problems, explicit attention to those concepts and algorithms is required. At the same time, we found no evidence that intervention on CALs transfers to WP outcomes, which corroborates prior research (e.g., Fuchs et al., 2011; Fuchs, Powell, et al., 2009). At the same time, we learned that a multitier support system focused on WPs improves WP performance of students at risk for CAL/WP-MD. However, in contrast to CAL-RTI, for which we saw no indication of transfer to WP skill, we did see evidence of transfer from WP-RTI to single-digit (but not double-digit) CAL problems. Given that competence with CALs and WPs is necessary for success in adulthood, findings suggest efficiency for WP-RTI over CAL-RTI, even as results indicate that instruction on WPs, which does not explicitly address CALs, does not transfer to improved performance on double-digit CAL problems that require regrouping.

Because the research base on prealgebraic reasoning is emerging, we were also interested in learning whether there were differences between CAL-RTI and WP-RTI for improving prealgebraic reasoning. WP-RTI improved the prealgebraic skill of students at risk for CAL/WP-MD significantly more than occurred in the CAL-RTI condition or in the CON group. This indicates that components of WP instruction, as designed in the present study to support the connection with prealgebraic thinking, contribute to successful prealgebraic reasoning. As algebraic reasoning is an integral part of mathematics standards at the elementary, middle, and high school grades (National Governors Association Center for Best Practices & Council of Chief State School Officers, 2010), and as many students are unprepared for algebra courses (NMAP, 2008), elementary school curricula need to emphasize algebraic reasoning. By using WP interventions focused on teaching the underlying structure of the problem, using equations to represent WP schema, embedding equal-sign instruction within WP intervention, and solving equations with a variable (e.g., x), the prealgebraic competence of students with CAL/WP-MD improves, and this may lead to better success with mathematics, especially algebra, in later grades.

In these ways, this study's WP and CAL multilevel support systems enhanced the mathematics performance of high-risk students—those with concurrent difficulty with CALs and WPs. These students, who are already performing substantially below grade-level expectations, must demonstrate improvement in WPs and CALs in order to meet ambitious mathematics standards in late elementary, middle, and high school. By design, to assess transfer across mathematics domains, our CAL-RTI focused exclusively on teaching CAL skill, with strong effects on CALs. Our WP-RTI, which focused exclusively on WPs, provided greater cognitive demands by challenging students to recognize WPs by type, write number sentences to represent WP schema, and solve equations with an unknown value. If teachers must choose between providing WP or CAL intervention, given constraints on mathematics intervention time in most schools, study results favor WP-RTI: Whereas CAL-RTI enhanced only CAL skill, WP-RTI simultaneously improved students' competence with WPs, single-digit CALs, and higher-level problems related to prealgebraic reasoning.

Implications for Practice

This study therefore suggests that teachers, when addressing the needs of students with co-occurring CAL and WP difficulty, must allocate specific instructional focus to WPs, rather than assuming that improvement in CALs will translate into improvement with WPs. Moreover, WP-RTI significantly improved prealgebraic reasoning more so than did CAL-RTI on prealgebraic reasoning. So, for teachers who have a limited window of time to improve the mathematics performance of students with CAL/WP-MD when early intervention is paramount (Dowker, 2005), WP-RTI provides students with an intervention package that improves a variety of mathematics skills rather than a single aspect of mathematics (i.e., CALs). Pirate Math, which is the WP-RTI program used in the present study, also includes explicit strategy instruction on CALs (note that CAL instruction was not provided in the present study's WP-RTI condition, to permit conclusions about transfer from WP instruction to CALs). Therefore, it appears that Pirate Math provides teachers with effective step-by-step methods for (a) teaching students how to set up and solve WPs, (b) embedding CAL strategy instruction practice within WPs, and (c) explicit and implicit practice on prealgebraic reasoning skills, such as understanding operations, setting up equations, and solving for unknowns within equations.

WP-RTI provides students with an intervention package that improves a variety of mathematics skills rather than a single aspect of mathematics.

Before automatically discounting CAL-RTI in favor on WP-RTI, however, we emphasize that all small-group tutoring students demonstrated low performance on CAL and WP at pretest. Although many students with MD experience difficulty in both areas, some students with MD experience difficulty with CAL without WP difficulty or vice versa (Fuchs, Fuchs, Stuebing, et al., 2008; Hanich, Jordan, Kaplan, & Dick, 2001). For students with CAL-MD only, teachers may want to focus on remediation of CAL deficits with CAL-RTI especially if students experience difficulty with double-digit CALs requiring regrouping. With CAL-RTI, however, students did not demonstrate significant improvement on prealgebraic reasoning. Teachers must provide prealgebraic preparation in some manner, and results show that such preparation can be effectively provided within the WP-RTI framework using Pirate Math.

To effectively implement either CAL-RTI or WP-RTI, school administrators must provide appropriate training for teachers or paraprofessionals. Implementation at both tiers must be conducted with high levels of fidelity to ensure students are receiving appropriate and effective instruction. In addition the school schedule must be designed to allocate 45 min of whole-class instruction twice each week and 30 min of small-group tutoring three times each week. This is a large commitment in terms of time and personnel, but if the mathematics performance of the majority of students improves (as in the present study) and the mathematics performance of students most at risk for MD significantly improves, then the commitment is necessary.

References

Alibali, M. W., Knuth, E. J., Hattikudur, S., McNeil, N. M., & Stephens, A. C. (2007). A longitudinal examination of middle school students' understanding of the equal sign and equivalent equations. *Mathematical Thinking and Learning*, *9*, 221–246. doi:10.1080/10986060701360902

Balacheff, N. (2001). Symbolic arithmetic vs. algebra: The core of a didactical dilemma. In R. Sutherland, T. Rojano, A. Bell, & R. C. Lins (Eds.), *Perspectives on school algebra* (pp. 249–260). Dordrecht, Netherlands: Kluwer. doi:10.1007/0-306-47223-6_14

Baroody, A. J., Bajwa, N. P., & Eiland, M. (2009). Why can't Johnny remember the basic facts? *Developmental Disabilities Research Reviews*, *15*, 69–79. doi:10.1002/ddrr.45

Baroody, A. J., & Ginsburg, H. P. (1983). The effects of instruction on children's understanding of the "equals" sign. *Elementary School Journal*, *84*, 199–212. doi:10.1086/461356

Blanton, M. L., & Kaput, J. J. (2005). Characterizing a classroom practice that promotes algebraic reasoning. *Journal for Research in Mathematics Education*, *36*, 412–446.

Bryant, D. P., Bryant, B. R., Gersten, R., Scammacca, N., & Chavez, M. M. (2008). Mathematics intervention for first- and second-grade students with mathematics difficulties: The effects of tier 2 intervention delivered as booster lessons. *Remedial and Special Education*, *29*, 20–32. doi:10.1177/0741932507309712

Bryant, D. P., Bryant, B. R., Gersten, R. M., Scammacca, N. N., Funk, C., Winter, A., . . . Pool, C. (2008). The effects of Tier 2 intervention on the mathematics performance of first-grade students who are at risk for mathematics difficulties. *Learning Disability Quarterly*, *31*, 47–63. doi:10.2307/20528817

Carraher, D. W., & Schliemann, A. D. (2002). Is everyday mathematics truly relevant to mathematics education? In J. Moshkovich & M. Brenner (Eds.), *Everyday mathematics: Monographs of the* Journal for Research in Mathematics Education (Vol. 11, pp. 131–153). doi:10.2307/749968

Chong, S. L., & Siegel, L. S. (2008). Stability of computational deficits in math learning disability from second through fifth grades. *Developmental Neuropsychology*, *33*, 300–317. doi:10.1080/87565640801982387

Clement, L. L., & Bernhard, J. Z. (2005). A problem-solving alternative to using key words. *Mathematics Teaching in the Middle School*, *10*, 360–365.

Connolly, A. J. (1998). *KeyMath–Revised*. Circle Pines, MN: American Guidance.

Cummins, D. D., Kintsch, W., Reusser, K., & Weimer, R. (1988). The role of understanding in solving word problems. *Cognitive Psychology*, *20*, 405–438. doi:10.1016/0010-0285(88)90011-4

Dowker, A. (2005). Early identification and intervention for students with mathematics difficulties. *Journal of Learning Disabilities*, *38*, 324–332. doi:10.1177/00222194050380040801

Falkner, K. P., Levi, L., & Carpenter, T. P. (1999). Children's understanding of equality: A foundation for algebra. *Teaching Children Mathematics*, *6*, 232–236.

Fuchs, L. S., Compton, D. L., Fuchs, D., Hollenbeck, K. N., Craddock, C., & Hamlett, C. L. (2008). Dynamic assessment of algebraic learning in predicting third graders' development of mathematical problem solving. *Journal of Educational Psychology*, *100*, 829–850. doi:10.1037/a0012657

Fuchs, L. S., Compton, D. L., Fuchs, D., Powell, S. R., Schumacher, R. F., Hamlett, C. L., . . . Vukovic, R. K. (2012). Contributions of domain-general cognitive resources and different forms of arithmetic development to pre-algebraic knowledge. *Developmental Psychology*, *48*, 1315–1326. doi:10.1037/a0027475

Fuchs, L. S., & Fuchs, D. (2002). Mathematical problem-solving profiles of students with mathematics disabilities with and without comorbid reading disabilities. *Journal of Learning Disabilities*, *35*, 563–573. doi:10.1177/00222194020350060701

Fuchs, L. S., Fuchs, D., Craddock, C., Hollenbeck, K. N., Hamlett, C. L., & Schatschneider, C. (2008). Effects of small-group tutoring with and without validated classroom instruction on at-risk students' math problem solving: Are two tiers of prevention better than one? *Journal of Educational Psychology*, *100*, 491–509. doi:10.1037/0022-0663.100.3.491

Fuchs, L. S., Fuchs, D., Hamlett, C. L., Phillips, N. B., Karns, K., & Dutka, S. (1997). Enhancing students' helping behavior during peer-mediated instruction with conceptual mathematical explanations. *Elementary School Journal*, *97*, 223–249. doi:10.1086/461863

Fuchs, L. S., Fuchs, D., Stuebing, K., Fletcher, J. M., Hamlett, C. L., & Lambert, W. E. (2008). Problem-solving and computation skills: Are they shared or distinct aspects of mathematical cognition? *Journal of Educational Psychology*, *100*, 30–47. doi:10.1037/0022-0663.100.1.30

Fuchs, L. S., Geary, D. C., Compton, D. L., Fuchs, D., Hamlett, C. L., & Bryant, J. V. (2010). The contributions of numerosity and domain-general abilities to school readiness. *Child Development*, *81*, 1520–1533. doi:10.1111/j.1467-8624.2010.01489.x

Fuchs, L. S., Geary, D. C., Compton, D. L., Fuchs, D., Schatschneider, C., Hamlett, C. L., . . . Changas, P. (2013). Effects of first-grade number knowledge tutoring with contrasting forms of practice. *Journal of Educational Psychology*, *105*, 58–77. doi:10.1037/a0030127

Fuchs, L. S., Hamlett, C. L., & Powell, S. R. (2003). *Second-Grade Calculations Battery*. Available from L. S. Fuchs, 228 Peabody, Vanderbilt University, Nashville, TN 37203.

Fuchs, L. S., Powell, S. R., Cirino, P. T., Schumacher, R. F., Marrin, S., Hamlett, C. L., . . . Changas, P. C. (in press). Does calculation or word-problem instruction provide a stronger route to pre-algebraic knowledge? *Journal of Educational Psychology*.

Fuchs, L. S., Powell, S. R., Hamlett, C. L., Fuchs, D., Cirino, P. T., & Fletcher, J. M. (2008). Remediating computational deficits at third grade: A randomized field trial. *Journal of Research on Educational Effectiveness*, *1*, 2–32. doi:10.1080/19345740701692449

Fuchs, L. S., Powell, S. R., Seethaler, P. M., Cirino, P. T., Fletcher, J. M., Fuchs, D., & Hamlett, C. L. (2011). The development of arithmetic and word-problem skill among students with mathematics disability. In J. M. Kauffman & D. P. Hallahan (Eds.), *Handbook of special education* (pp. 434–444). New York, NY: Taylor & Francis.

Fuchs, L. S., Powell, S. R., Seethaler, P. M., Cirino, P. T., Fletcher, J. M., Fuchs, D., . . . Zumeta, R. O. (2009). Remediating number combination and word problem deficits among students with mathematics difficulties: A randomized control trial. *Journal of Educational Psychology*, *101*, 561–576. doi:10.1037/a0014701

Fuchs, L. S., Seethaler, P. M., & Powell, S. R. (2009). *Test of Pre-Algebraic Reasoning*. Available from L. S. Fuchs, 228 Peabody, Vanderbilt University, Nashville, TN 37203.

Fuchs, L. S., Seethaler, P. M., Powell, S. R., Fuchs, D., Hamlett, C. L., & Fletcher, J. M. (2008). Effects of preventative tutoring on the mathematical problem solving of third-grade students with math and reading difficulties. *Exceptional Children*, *74*, 155–173.

Fuchs, L. S., Zumeta, R. O., Schumacher, R. F., Powell, S. R., Seethaler, P. M., Hamlett, C. L., & Fuchs, D. (2010). The effects of schema-broadening instruction on second graders' word-problem performance and their ability to represent word problems with algebraic equations: A randomized control study. *Elementary School Journal*, *110*, 440–463. doi:10.1086/651191

Fuson, K. C., & Kwon, Y. (1992). Korean children's single-digit addition and subtraction: Numbers structured by ten. *Journal for Research in Mathematics Education*, *23*, 148–165. doi:10.2307/749498

Geary, D. C., Boykin, A. W., Embretson, S., Reyna, V., Siegler, R., Berch, D. B., & Graban, J. (2008). Report of the Task Group on Learning Processes. In National Mathematics Advisory Panel, *Reports of the task groups and subcommittees* (pp. 4-i–4-221). Washington, DC: U.S. Department of Education. Retrieved from http://www2.ed.gov/about/bdscomm/list/mathpanel/report/learning-processes.pdf

Gersten, R., Beckmann, S., Clarke, B., Foegen, A., March, L., Star, J. R., & Witzel, B. (2009). *Assisting students struggling with mathematics: Response to intervention (RtI) for elementary and middle schools* (NCEE 2009-4060). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. Retrieved from http://ies.ed.gov/ncee/wwc/publications/practiceguides/

Greenes, C., Larson, M., Leiva, M. A., Shaw, J. M., Stiff, L., Vogeli, B. R., & Yeatts, K. (2005). *Houghton Mifflin math*. Boston, MA: Houghton Mifflin.

Groen, G., & Resnick, L. B. (1977). Can preschool children invent addition algorithms? *Journal of Educational Psychology*, *69*, 645–652. doi:10.1037/0022-0663.69.6.645

Hanich, L. B, Jordan, N. C., Kaplan, D., & Dick, J. (2001). Performance across different areas of mathematics cognition in children with learning difficulties. *Journal of Educational Psychology*, *93*, 615–626. doi:10.1037//0022-0663.93.3.615

Hoover, H. D., Hieronymous, A. N., Dunbar, S. B., & Frisbie, D. A. (1993). *Iowa Test of Basic Skills, Form K*. Itasca, IL: Riverside.

Jacobs, V. R., Franke, M. L., Carpenter, T. P., Levi, L., & Battey, D. (2007). Professional development focused on children's algebraic reasoning in elementary school. *Journal for Research in Mathematics Education*, *38*, 258–288.

Jitendra, A. K. (2007). *Solving math word problems: Teaching students with learning disabilities using schema-based instruction*. Austin, TX: Pro-Ed.

Jitendra, A. K., Hoff, K., & Beck, M. M. (1999). Teaching middle school students with learning disabilities to solve word problems using a schema-based approach. *Remedial and Special Education*, *20*, 50–64. doi:10.1177/074193259902000108

Jordan, N. C., & Hanich, L. (2000). Mathematical thinking in second-grade children with different forms of LD. *Journal of Learning Disabilities*, *33*, 567–578. doi:10.1177/002221940003300605

Kieran, C. (1990). Cognitive processes involved in learning school algebra. In P. Nesher & J. Kilpatrick (Eds.), *Mathematics and cognition: A research synthesis by the International Group for the Psychology of Mathematics Education* (pp. 96–112). New York, NY: Cambridge University Press. doi:10.1017/CBO9781139013499.007

Kieran, C. (1992). The learning and teaching of school algebra. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 390–419). Reston, VA: National Council of Teachers of Mathematics.

Kintsch, W., & Greeno, J. G. (1985). Understanding and solving word arithmetic problems. *Psychological Review*, *92*, 109–129. doi:10.1037/0033-295X.92.1.109

Knuth, E. J., Stephens, A. C., McNeil, N. M., & Alibali, M. W. (2006). Does understanding the equal sign matter? Evidence from solving equations. *Journal for Research in Mathematics Education*, *37*, 297–312.

Lee, K., Ng, S. F., Bull, R., Pe, M. L., & Ho, R. H. M. (2011). Are patterns important? An investigation of the relationships between proficiency in patterns, computation, executive functioning, and algebraic word problems. *Journal of Educational Psychology*, *103*, 269–281. doi:10.1037/a0023068

LeFevre, J. A., & Morris, J. (1999). More on the relation between division and multiplication in simple arithmetic: Evidence for mediation of division solutions via multiplication. *Memory & Cognition*, *27*, 803–812. doi:10.3758/BF0 3198533

Lembke, E. S., Garman, C., Deno, S. L., & Stecker, P. M. (2010). One elementary school's implementation of response to intervention (RTI). *Reading and Writing Quarterly*, *26*, 361–373. doi:10.1080/10573569.2010.500266

Linchevski, L. (2001). Operating on the unknowns: What does it really mean? In M. van den Heuvel-Panhuizen (Ed.), *Proceedings of the 25th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 1, pp. 141–144). Utrecht, Netherlands: PME Proceedings.

Linchevski, L., & Herscovics, N. (1996). Crossing the cognitive gap between arithmetic and algebra: Operating on the unknown in the context of equations. *Educational Studies in Mathematics*, *30*, 39–65. doi:10.1007/BF00163752

Luke, D. A. (2004). *Multilevel modeling (quantitative applications in the social sciences)*. Thousand Oaks, CA: SAGE.

Ma, X., & Ma, L. (2008). Using multilevel modeling to investigate school effects. In A. A. O'Connell & D. B. McCoach (Eds.), *Multilevel modeling of educational data* (pp. 59–110). Charlotte, NC: Information Age.

McNeil, N. M., & Alibali, M. W. (2005). Why won't you change your mind? Knowledge of operational patterns hinders learning and performance on equations. *Child Development*, *76*, 883–899. doi:10.1111/j.1467-8624.2005.00884.x

Mellard, D. F., Stern, A., & Woods, K. (2011). RTI school-based practices and evidence-based models. *Focus on Exceptional Children*, *43*(6), 1–15.

Murnane, R. J., Willett, J. B., Braatz, M. J., & Duhaldeborde, Y. (2001). Do different dimensions of male high school students' skills predict labor market success a decade later? Evidence from the NLSY. *Economics of Education Review*, *20*, 311–320. doi:10.1016/S0272-7757(00)00056-X

Nathan, M. J., Kintsch, W., & Young, E. (1992). A theory of algebra-word-problem comprehension and its implications for the design of learning environments. *Cognitive and Instruction*, *9*, 392–389. doi:10.1207/s1532690xci0904_2

National Center for Education Statistics. (2008). *Digest of education statistics*. Washington, DC: U.S. Department of Education. Retrieved from http://nces.ed.gov/pubs2009/2009020.pdf.

National Governors Association Center for Best Practices & Council of Chief State School Officers. (2010). *Common core state standards mathematics*. Washington, DC: Authors.

National Mathematics Advisory Panel. (2008). *Final report*. Washington, DC: U.S. Department of Education.

Parsons, S., & Bynner, J. (1997). Numeracy and employment. *Education and Training*, *39*, 43–51. doi:10.1108/00400919710164125

Peugh, J. L. (2010). A practical guide to multilevel modeling. *Journal of School Psychology*, *48*, 85–112. doi:10.1016/j.jsp.2009.09.002

Pillay, H., Wilss, L., & Boulton-Lewis, G. (1998). Sequential development of algebra knowledge: A cognitive analysis. *Mathematics Education Research Journal*, *10*, 87–102. doi:10.1007/BF03217344

Powell, S. R. (2012). Equations and the equal sign in elementary mathematics textbooks. *Elementary School Journal*, *112*, 627–648. doi:10.1086/665009

Powell, S. R., & Fuchs, L. S. (2010). Contribution of equal-sign instruction beyond word-problem tutoring for third-grade students with mathematics difficulty. *Journal of Educational Psychology*, *102*, 381–394. doi:10.1037/a0018447

Powell, S. R., & Fuchs, L. S. (2014). *Does early algebraic reasoning differ as a function of students' difficulty with calculations versus word problems?* Manuscript submitted for publication.

Psychological Corporation. (1999). *Wechsler Abbreviated Scale of Intelligence (WASI) manual*. San Antonio, TX: Author.

RAND Mathematics Study Panel. (2003). *Mathematical proficiency for all students: Toward a strategic research and development program in mathematics education* (No. 08330331X). Santa Monica, CA: RAND.

Riley, M. S., & Greeno, J. G. (1988). Developmental analysis of understanding language about quantities and of solving problems. *Cognition and Instruction*, *5*, 49–101.

Rittle-Johnson, B., & Alibali, M. W. (1999). Conceptual and procedural knowledge of mathematics: Does one lead to the other? *Journal of Educational Psychology*, *91*, 175–189. doi:10.1037/0022-0663.91.1.175

Sfard, A. (1991). On the dual nature of mathematical conceptions: Reflections on processes and objects as different sides of the same coin. *Educational Studies in Mathematics*, *22*, 1–16. doi:10.1007/BF00302715

Sfard, A., & Linchevski, L. (1994). The gains and the pitfalls of reification: The case of algebra. *Educational Studies in Mathematics*, *26*, 191–228. doi:10.1007/BF01273663

Siegler, R. S., & Shrager, J. (1984). Strategy choice in addition and subtraction: How do children know what to do? In C. Sophian (Ed.), *Origins of cognitive skills* (pp. 229–293). Hillsdale, NJ: Lawrence Erlbaum.

Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. Oxford, England: Oxford University Press.

Sullivan, L. M., Dukes, K. A., & Losina, E. (1999). Tutorial in biostatistics: An introduction to hierarchical linear modeling. *Statistics in Medicine*, *18*, 855–888.

Swanson, H. L. (2006). Cross-sectional and incremental changes in working memory and mathematical problem solving. *Journal of Educational Psychology*, *98*, 265–281. doi:10.1037/0022-0663.98.2.265

Tolar, T. D., Lederberg, A. R., & Fletcher, J. M. (2009). A structural model of algebra achievement: Computational fluency and spatial visualization as mediators of the effect of working memory on algebra achievement. *Educational Psychology*, *29*, 239–266. doi:10.1080/01443410802708903

Tournaki, N., Bae, Y. S., & Kerekes, J. (2008). Rekenrek: A manipulative used to teach addition and subtraction to students with learning disabilities. *Learning Disabilities: A Contemporary Journal*, *6*(2), 41–59.

Vaughn, S., & Fuchs, L. S. (2003). Redefining learning disabilities as inadequate response to instruction: The promise and potential problems. *Learning Disabilities Research & Practice*, *18*, 137–146. doi:10.1111/1540-5826.00070

Wei, X., Lenz, K. B., & Blackorby, J. (2012). Math growth trajectories of students with disabilities: Disability category, gender, racial, and socioeconomic status differences from ages 7 to 17. *Remedial and Special Education*, *34*, 154–165. doi:10.1177/0741932512448253

Wilkinson, G. S. (1993). *Wide Range Achievement Test 3*. Wilmington, DE: Wide Range.

Authors' Note

This research was supported by Award No. R01 HD059179 and Core Grant No. HD15052 from the Eunice Kennedy Shriver National Institute of Child Health and Human Development to Vanderbilt University. The content is solely the responsibility of the authors and does not necessarily represent the official views of the Eunice Kennedy Shriver National Institute of Child Health and Human Development or the National Institutes of Health.

Manuscript received November 2013; accepted March 2014.